

METHODS

A generalized approach for Bayesian Gaussian graphical models

Vithor Rosa Franco^{1*}, Guilherme W. F. Barros², & Marcos Jiménez³

Received: October 1, 2023 | **Accepted:** August 9, 2024 | **Published:** August 16, 2024 | **Edited by:** Hudson Golino

¹ Postgraduate Program of Psychology, São Francisco University; ² Umeå School of Business, Economics and Statistics (USBE), Umeå University; ³ Department of Social Psychology and Methodology, Universidad Autónoma de Madrid.

*Please address correspondence to Vithor Rosa Franco, vithorfranco@gmail.com, Universidade São Francisco - Campus de Campinas. Rua Waldemar César da Silveira, 105, Jardim Cura D'Ars. 13045510 - Campinas, SP - Brazil. This article is published under the Creative Commons BY 4.0 license. Users are allowed to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

Bayesian Gaussian Graphical Models (BGGMs) are tools of growing popularity and interest in network psychometrics and probabilistic graphical modeling. However, some of the existing models are derived from different modeling principles that do not easily allow for extensions and combinations into new models. More specifically, the implementation of some models may not be flexible enough to test different priors or likelihoods. In this paper, we present a new approach to BGGMs that overcomes this limitation by allowing for the estimation of regularized partial correlations between any type of variables while also having an intuitive approach on how to decide about the priors. Our approach is based on using a transformation of the lower diagonal values of the Cholesky (or LDL) decomposition matrix as the parameters of the models, which can receive any zero-centered symmetric distribution as a prior, as well as to include moderators. We have developed the *gbggm* R package to implement some models based on this approach, and the potentials of the approach are demonstrated with a toy simulation and an empirical example. This new approach expands the range of applications and enhances the flexibility of BGGMs, making them more useful in a variety of contexts.

Keywords: Bayesian analysis, probabilistic graphical modeling, psychometrics, regularization, computational statistics

1. INTRODUCTION

Probabilistic graphical models (Lauritzen, 1996) have garnered increasing interest in network psychometrics, finding applications in diverse fields such as mental health (Hallquist et al., 2021), intelligence (Van Der Maas et al., 2017), and personality (Costantini et al., 2019). Many of these studies assume multivariate normality and employ Maximum Likelihood methods (Isvoranu & Epskamp, 2023) to fit Gaussian graphical models (GGMs). Some researchers have explored the use of Bayesian approaches (e.g., Williams & Mulder, 2020a) for GGM estimation. Despite the advantages of Bayesian models, some limitations of existing Bayesian GGM methods may hinder the development of the field. For instance, currently, different data types (e.g., binary or ordinal vs. continuous data) require the development of new models and sampling schemes (e.g., Marsman et al., 2022). However, Bayesian GGMs can be formulated with more general and flexible model specifications. This study aims to introduce a general approach for regularized Bayesian Gaussian graphical models (GBGGMs) and to illustrate the performance of proposed models through a toy simulation and an empirical example.

1.1 Regularized Gaussian Graphical Models (GGMs)

A graph can be defined as a pair $G = (V, E)$, where V represents a set of nodes and E the edges connecting such nodes. E is by itself a subset of the set $V \times V$ which consists of pairs of distinct nodes, resulting in G having no cases with multiple edges or loops (Lauritzen, 1996). In a Gaussian Graphical Model (GGM), each node of the graph represents a normally distributed variable, and the set of observed variables in the model follow a multivariate normal distribution. Each edge in a GGM represents the partial correlation between each node, as the variables are conditionally dependent given the remaining variables in the network (Epskamp, Waldorp, et al., 2018). These partial correlations are obtained from the

concentration matrix Σ^{-1} , which is the inverse of the covariance matrix Σ . From the elements p of the concentration matrix, the partial correlation for a pair of variables i and j , ω_{ij} , is defined as:

$$\omega_{ij} = -\frac{p_{ij}}{\sqrt{p_{ii}p_{jj}}} \quad (1)$$

Compared to Pearson correlation, partial correlations between two variables have the advantage of being adjusted for the influence of other variables, thus avoiding spurious correlations. Additionally, under specific circumstances, partial correlations can indicate causal relationships between variables (Andersson et al., 1997; Epskamp, Maris, et al., 2018). In the graphical representation of a GGM, the absence of an edge between nodes indicates the absence of a relationship, whether marginal or conditional, between a pair of variables.

The application of network methods in psychological research has gained momentum over the past five years. According to Isvoranu and Epskamp (2023), several factors contribute to this trend, including the need for more intuitive and visual representations of psychological constructs and their interrelations, the exploration of potential causal relationships in correlational studies, and efforts to address the limitations of reflective theories of mental constructs. Although the use of network methods in psychology has recently surged, its roots trace back to the 1930s when Moreno (1934) first proposed their use in studying social networks or sociograms. Moreno's pioneering work primarily focused on the social connections individuals establish. Consequently, networks predominantly depicted relationships between individuals rather than observed variables (Wasserman & Faust, 1994). In contrast, contemporary network psychometrics (Epskamp, Maris, et al., 2018) employs graphical models where nodes represent variables and edges represent statistical relationships between them (Isvoranu & Epskamp, 2023). This approach bears similarities to Structural Equation Models (SEM) but

from a more exploratory perspective (Epskamp, 2020).

A crucial aspect of Gaussian Graphical Models (GGMs) involves identifying partial correlations that can be reliably inferred as exactly zero (Williams, 2020). Many estimated relationships in GGMs may be spurious correlations, warranting a method to address this issue. Regularization methods, such as the graphical LASSO (or “glasso”), offer a solution (Friedman et al., 2008). The glasso technique involves applying a modified LASSO regression to each variable, which effectively forces small partial correlation coefficients to zero. Similar to LASSO regression, the glasso method utilizes a tuning parameter to control the degree of sparsity in the resulting graph. Lower values of this parameter yield denser graphs, while higher values produce sparser ones. Additionally, glasso can be combined with the Extended Bayesian Information Criterion (EBIC) fit statistic for model selection. This combined approach has proven to be effective in determining an optimal value for the tuning parameter (Isvoranu & Epskamp, 2023).

Despite the widespread use of glasso, Williams (2020) has argued that it: (i) leads to overshrinkage when the sample size is smaller than the number of variables; (ii) leads to overshrinkage when a set of predictors is correlated; (iii) has a higher prediction error than the ridge when the predictors are highly correlated; (iv) can lead to overshrinkage of large coefficients; and (v) does not perform well if the true data generating process is not known. Due to these limitations, alternative regularization methods for GGMs have been proposed by various authors (see Williams (2020) for a survey of alternative Maximum Likelihood methods). This scenario highlights the necessity of developing new regularization methods for GGMs, given that a specific method may fail in specific contexts. Despite of growing interest in Bayesian approaches for GGMs, they still lag behind in terms of diversity compared to Maximum Likelihood-based methods.

1.2 The Standard Bayesian Estimation of GGMs: Williams (2021)

In response to the limitations of regularization methods within the Maximum Likelihood approach, several authors have proposed Bayesian alternatives (e.g., Byrd et al., 2021; Li et al., 2019; Liu & Martin, 2019; Marlin & Murphy, 2009; Marsman & Haslbeck, 2023; A. Mohammadi & Wit, 2015; R. Mohammadi et al., 2021; Park et al., 2022; Sagar et al., 2024; Wang, 2012). In the context of network psychometrics, probably the most well-known approach is one proposed by Williams (2021). The author proposes that estimation follows a conventional Bayesian approach used for modeling multivariate normal variables by sampling the covariance matrix from an inverse Wishart prior (Sun & Berger, 2007). The Wishart distribution involves two hyperparameters: the base covariance matrix V and the degree of freedom κ , where κ is equal to or greater than the number of variables. This estimation approach aligns with the estimation of regularized Bayesian Gaussian graphical models (BGGMs), wherein covariance or correlation values are shrunk, or regularized, toward zero. This is achieved by setting V as a diagonal (or identity) matrix, with κ determining the degree of regularization in the model; higher values indicate stronger regularization. To discern the presence or absence of edges, post-hoc statistics based on the posterior distribution of edges are employed. A straightforward strategy entails examining whether zero is encompassed within the posterior distribution of estimated partial correlations. If so, the corresponding edge should be omitted from the final estimate of the GGM.

While the approach proposed by Williams (2021) has notable strengths, we argue that a generalized Bayesian GGM approach should possess certain key capabilities: (i) the ability to model correlations across diverse data types; (ii) the provision of parameter estimates indicating the probability of edge inclusion; and (iii) flexibility and ease of implementation with general-purpose optimization algorithms. For the best

of our knowledge, these desiderata are not implemented concomitantly in any current Bayesian GGM method. In regards to William's approach, from this standpoint, we identify two primary limitations. Firstly, it is tailored specifically for continuous data. Although similar methods exist that employ extensions of the Gibbs sampling scheme (e.g., Albert & Chib, 1993; Hoff, 2007; Talhouk et al., 2012), and are incorporated into the BGGM R package (Williams & Mulder, 2020b), these approaches are heavily reliant on their particular sampling schemes, limiting their adaptability to alternative estimation methods such as different Markov Chain Monte Carlo (MCMC) algorithms or variational Bayes and Laplace Approximation. Secondly, Williams's approach does not directly estimate the presence or absence of edges during the estimation process. Instead, this information is derived post-estimation from the posterior distribution of parameters using procedures like the Savage-Dickey ratio (Williams & Mulder, 2020a).

1.3 Cholesky Decomposition of Correlation Matrices

One main challenge with the Bayesian estimation of GGMs is the sampling of positive semi-definite matrices. In the case of Markov Chain Monte Carlo (MCMC) estimation (Hanada & Matsuura, 2022), a sampler based on the Wishart distribution can be developed, with the base covariance matrix V set as the current sample of the covariance matrix, and the degrees of freedom κ representing the "precision" of the sampler: larger values of κ will make the sampler draw covariance matrices closer to V . One limitation of this approach is that it involves sampling the entire matrix simultaneously, which limit the precision of the posterior distributions of the estimated covariances to be equal. Also, the sampling of the entire matrix at once does not allow to fix a subset of parameters at known (or assumed to be known) values.

Additionally, in the scenario of point estimation using deterministic optimization methods like

gradient descent, another challenge arises: all (co)variances must be simultaneously constrained to ensure the resulting estimated matrix is positive semi-definite. The Generalized Bayesian Gaussian Graphical Model (GBGGM) approach proposed in this paper, which is suitable both for MCMC as well as deterministic optimization, has as its initial step the use of the Cholesky decomposition approach for the estimation of the correlation parameters (Archakov et al., 2024; Archakov & Hansen, 2021; Bhat & Mondal, 2021).

The utilization of Cholesky decomposition for estimating correlation matrices is well-documented in the literature (the interested reader is referred to Chen & Leng, 2015; Ghosh et al., 2021; Roverato, 2000; Verzele, 2010; Ye et al., 2020). In this section, we outline the fundamental steps for its Bayesian estimation. The Cholesky decomposition of matrix A is defined as:

$$A = LL^T \quad (2)$$

where L is a lower triangular matrix and L^T is its transpose. This decomposition guarantees that A will be a positive semi-definite matrix, as long as the diagonal elements of L are equal or larger than 0. For the estimation of a correlation matrix, the lower-diagonal elements of L are treated separately and can be shifted (in the case of deterministic optimization), or sampled (in the case of the MCMC estimation), independently. However, to guarantee that A will be a correlation matrix R , some additional steps are necessary. More specifically, starting with a lower triangular matrix C , where all the diagonal elements are equal to 1 and the lower-diagonal elements, or the "weights", are real, L is calculated with:

$$L = (C^T E)^T \quad (3)$$

where E is a diagonal matrix with its diagonal values equal to e_i , for $i = 1, 2, 3, \dots, k$ is the order of E , and e_i is equal to the inverse of the Euclidean norm of the i -th row of C . The estimation of correlation matrices with the Cholesky

decomposition, then, can be conducted on the elements of C , also called in this paper as “weights”, which are used to estimate the correlation and partial correlation matrices (a similar approach is also found in Lewandowski et al., 2009). To estimate a covariance matrix, this approach can be extended by using the LDL decomposition (Krishnamoorthy & Menon, 2013), which is defined as:

$$A = LDL^T \quad (4)$$

where D is a positive diagonal matrix and L is the same as before.

1.4 A General Approach for Bayesian Estimation of GGMs

In our generalized approach to BGGMs, the modeling and estimation process includes an additional step. If we define Q as the inverse of the lower-diagonal matrix L , $Q = L^{-1}$, and considering that the partial correlation matrix is derived from the inverse of the correlation matrix, we can exploit the fact (Krishnamoorthy & Menon, 2013) that

$$A^{-1} = Q^T Q \quad (5)$$

to set a prior to the lower-diagonal elements of Q , which will correspond to the weights of the partial correlation matrix. Then, to estimate a regularized partial correlation matrix, one can use shrinkage priors on the $p = k(k-1)/2$ lower-diagonal elements (i.e., q_j) of Q , as similarly performed in Bayesian penalized regression (van Erp, 2020). In summary, we set the same symmetric distribution as the prior for all the lower-diagonal elements of Q , where this prior has to have its location parameter equal to 0 and its scale parameter serving as the inverse of the strength of the regularization (i.e., amount of shrinkage). It is also possible to use distributions with additional parameters, as long as there are at least a location and a scaling parameter.

From this, the general representation of the

approach is the following:

$$q_j \sim P(0, \boldsymbol{\theta}), \text{ for } j = 1, \dots, p, \quad (6)$$

In this case, we are using $\boldsymbol{\theta}$ to represent the set of additional parameters (i.e., other possible parameters that are not a location parameter, which is always set to 0) that a distribution may have, including, at least, a scaling parameter. Therefore, for instance, ridge regularization (van Erp et al., 2019) is achieved by setting the prior $P(0, \boldsymbol{\theta})$ to be the Normal distribution with a mean of 0 and standard deviation of λ . More specifically, we have

$$q_j \sim N\left(0, \frac{1}{\lambda}\right), \text{ for } j = 1, \dots, p, \quad (7)$$

$$\lambda \sim \text{Cauchy}^+(0, 1)$$

where λ is the shrinkage parameter, p is the number of parameters, and Cauchy^+ is the half-Cauchy distribution. Bayesian LASSO regularization (van Erp et al., 2019), in contrast, is achieved with

$$q_j \sim \text{Laplace}\left(0, \frac{1}{\lambda}\right), \text{ for } j = 1, \dots, p, \quad (8)$$

$$\lambda \sim \text{Cauchy}^+(0, 1)$$

where *Laplace* stands for the Laplace, or double-exponential, distribution.

There are several generalizations and extensions of the ridge and LASSO methods in the Bayesian literature (van Erp et al., 2019). In general, any symmetric distribution centered at zero (Déniz, 2021) can be used as the prior for the elements of Q . For instance, one could use the logistic or the Cauchy distributions instead of the normal or the Laplace distributions. One could also add an extra parameter to the model and then set the prior to be equal to the three-parameter t distribution (Lange et al., 1989), a heavy-tailed version of the ridge, or to the double lomax distribution (Bindu & Sangita, 2015), a heavy-tailed version of the LASSO. Further still, it is possible to set the prior as less “well-behaved” distributions, such as the generalized normal (Nadarajah, 2005), the skew-normal (Arellano-Valle & Azzalini, 2006), or the Normal-

exponential-gamma (Scheipl & Kneib, 2009) distributions. However, caution should be taken when choosing these priors, as the shrinkage properties of these alternatives may not be appropriate for the objective at hand. As argued before, in regards to the LASSO regularization, the performance of different regularization methods can be sub-optimal.

1.5 Defining the Likelihood of the GBGGM

After setting the prior for the elements of Q , the next step is to define the likelihood function that represents the data. For continuous data, the likelihood function is typically set as the multivariate normal distribution. For example, when dealing exclusively with continuous data and aiming to implement a correlation model with Bayesian LASSO regularization, the complete model can be specified as follows

$$\begin{aligned} q_j &\sim \text{Laplace}\left(0, \frac{1}{\lambda}\right), \text{ for } j = 1, \dots, p, \\ \lambda &\sim \text{Cauchy}^+(0, 1) \\ \hat{L} &\leftarrow (\hat{C}^T \hat{E})^T \\ \hat{R} &\leftarrow \hat{L} \hat{L}^T \\ X &\sim \text{MVN}(\mathbf{0}, \hat{R}) \end{aligned} \quad (9)$$

In psychometrics, a critical modeling challenge involves estimating polychoric correlations (Jöreskog, 1994). Polychoric correlations posit that the observed monotonic relationships among ordered categorical variables stem from discretizing linearly related Gaussian variables with unobserved thresholds. This assumption permits estimation of joint response proportions for ordered categorical variables based on the probability function of a multivariate normal distribution. However, the full Maximum Likelihood method is often impractical due to the exponential growth in joint responses with the number of variables and response categories (Olsson, 1979). An alternative feasible method within the Maximum Likelihood framework is to estimate polychoric correlations from bivariate normal marginal distributions given thresholds (Jöreskog, 1994). Yet, this method often yields a correlation matrix that is not positive definite. In contrast, our proposed generalized approach ensures the

estimated correlation matrix is positive definite while benefiting from the computational efficiency of calculating likelihood from bivariate normal marginal distributions. Additionally, our method enables simultaneous estimation of thresholds without necessitating a Bayesian two-step alternative (Lee & Poon, 1987).

Formally, the logarithm of the likelihood for polychoric correlations between each pair of variables i and j is defined as $\ln L_{ij} = N \sum_{a=1}^{M_{ij}} p_{ija} \ln \pi_{ija}(r_{ij}, \tau_{ija})$. Here, N is the sample size and M_{ij} represents the total number of possible combinations of response categories for variables i and j . For instance, if variable i has 3 response categories and j has 4 response categories, then $M_{ij} = 12$. p_{ija} denotes the proportion of responses in the combination a of response categories for variables i and j , and $\pi_{ija}(r_{ij}, \tau_{ija})$ is the cumulative probability of the bivariate normal distribution with correlation r_{ij} and threshold parameter τ_{ija} . For simplicity, we define $\text{PolyBVN}(\hat{R}, \boldsymbol{\tau})$ as the polychoric likelihood based on the bivariate normal distribution, where \hat{R} is the estimated correlation matrix and $\boldsymbol{\tau}$ is the vector of threshold parameters for all the items. The final model can be specified as follows:

$$\begin{aligned} q_j &\sim \text{Laplace}\left(0, \frac{1}{\lambda}\right), \text{ for } j = 1, \dots, p, \\ \lambda &\sim \text{Cauchy}^+(0, 1) \\ \hat{L} &\leftarrow (\hat{C}^T \hat{E})^T \\ \hat{R} &\leftarrow \hat{L} \hat{L}^T \\ X &\sim \text{PolyBVN}(\hat{R}, \boldsymbol{\tau}) \\ \boldsymbol{\tau} &\sim N(0, 1) \end{aligned} \quad (10)$$

1.6 Estimating the Sparsity of the Network

As a key feature of our proposed generalized Bayesian Gaussian Graphical Model method, we aimed to incorporate the estimation of network sparsity as part of the model per se. In previous Bayesian GGM methods found in the literature, sparsity of the correlation matrix was typically estimated through post-hoc procedures conducted after model fitting. In contrast, we propose utilizing a mixture probability model akin to the approach suggested by

Gorbach et al. (2020). The main idea with this approach is that the off-diagonal elements of the partial correlation matrix implied by the model form groups of ‘connected’ and ‘disconnected’ nodes. To achieve this, the extension of the model requires the calculation of the implied partial correlation matrix $\hat{\Omega}$ and some additional parameters.

While complex models like the one proposed by Gorbach et al. (2020) may enhance estimation quality, we advocate for a simpler approach. We exemplify our proposition with the same example as before: imagine we are dealing with continuous data and seeking to implement a correlation model with Bayesian LASSO regularization. If one desires to estimate a sparse partial correlation matrix, the complete model can be specified as follows

$$\begin{aligned}
 q_j &\sim \text{Laplace}\left(0, \frac{1}{\lambda}\right), \text{ for } j = 1, \dots, p, \\
 \lambda &\sim \text{Cauchy}^+(0, 1) \\
 \hat{L} &\leftarrow (\hat{C}^T \hat{E})^T \\
 \hat{R} &\leftarrow \hat{L} \hat{L}^T \\
 X &\sim \text{MVN}(\mathbf{0}, \hat{R}) \\
 \delta_j &\leftarrow H\phi(|\theta \cdot F(\omega_j)|, \sigma), \text{ for } j = 1, \dots, p, \\
 \theta &\sim \text{Cauchy}^+(0, 1) \\
 \sigma &\sim \text{Cauchy}^+(0, 1) \\
 F(\omega_j) &\sim (1 - \delta_j)N(0, \gamma) + \delta_j \text{Bernoulli}(\delta_j), \text{ for } j = 1, \dots, p, \\
 \gamma &\sim \text{Cauchy}^+(0, 1)
 \end{aligned} \tag{11}$$

The initial step in the extension proposed within this model involves computing δ_j , representing the probability of the edge j being present. This probability is calculated as a scaled transformation of the partial correlation ω_j . Specifically, $F(\cdot)$ is the Fisher transformation (i.e., the arc-tangent function) and θ is a scaling parameter. $H\phi$ is the cumulative function of the half-normal distribution, which depends on the parameter σ . Under this setup, larger values of θ increase the slope of the relation between the magnitude of the partial correlation and the probability of presence of the edge. Larger values of σ diminish the maximum possible probability of edge presence across all edges. Subsequently, the likelihood (or, probably more adequately, a hyperprior) of the Fisher-

transformed partial correlations $F(\omega_j)$ is calculated using a mixture distribution. The mixture comprises a normal density, with mean 0 and standard deviation defined by the γ parameter, weighted by $(1 - \delta_j)$, and a Bernoulli mass, with the probability parameter equal to δ_j , weighted by δ_j . The scaling parameters θ , σ , and γ are assigned half-Cauchy priors.

This extended version of the basic generalized Bayesian Gaussian graphical model introduces an alternative method for determining whether a node should be included in the final estimation of the partial correlation matrix. Notably, this approach eliminates the need for post-estimation methods to deduce the presence or absence of an edge from the posterior distribution of parameters. Instead, this model enables the use of point estimation, rather than full posterior distribution estimation, to make conclusions regarding edge presence. While the richness of Bayesian inference typically arises from the posterior distribution, in certain applications, the Maximum A Posteriori (MAP) estimate may suffice. Also, the post-estimation methods can still be used with the resulting estimates, if the full posterior distribution is estimated. Therefore, our model adds flexibility, but also additional information that can be used in the decision process about the true network that generates the data.

In summary, our suggestion for a generalized approach in estimation of BGGMs consists of the following steps:

1. Sample the elements of C .
2. With C , estimate L , Q , R , and Ω .
3. Choose a prior distribution for the lower-diagonal elements q_j of Q .
4. Choose any extra necessary shrinkage parameters according to sparsity constraints.
5. Calculate the likelihood of your data.

2. ILLUSTRATION WITH A TOY SIMULATION

In this section, we conduct a simple toy simulation to provide an initial assessment of the feasibility of our approach. To achieve this, we simulated a network with 7 nodes (equating to 21 correlations) based on a random Directed Acyclic Graph [DAG; Pearl (2009)], using the “ordered” method from the bnlearn R package (Scutari & Silander, 2024). From the random DAG, we derived its moral graph (i.e., the undirected graph that represents the equivalence class of the DAG) and, for the included edges, we generated random values using a random distribution with lower bound of .30 and upper bound of .70. If the values did not allow the transformation of this base matrix (i.e., the partial correlation matrix) to a correlation matrix, we repeated the sampling procedure until this transformation was feasible. The resulting partial correlation matrix is presented in Table 1.

The simulated observations for the 7 variables were sampled from a multivariate normal distribution with 200 cases. The means of the variables were set at 0, and the base correlation matrix was derived from the partial correlation

matrix shown in Table 1. In the subsequent step, the variables were dichotomized using thresholds defined by an equally spaced sequence ranging from -1.5 to 1.5 ; i.e., the threshold for the first variable was -1.5 , for the second variable it was set to -1.0 , and so forth.

We fitted the model proposed by Williams (2021) along with eight variants of the sparse Generalized Bayesian Gaussian Graphical Model (GBGGM) to the toy simulation data. The generalized models were implemented using the gbggm R package (Franco & Jimenez, 2023). The sole distinction among these models lay in the prior specification of the q_j weights. Specifically, we implemented the following prior distributions: normal; Laplace; logistic; Cauchy; three-parameter t ; double lomax; hyperbolic secant; and Normal-Exponential-Gamma. The likelihood was always calculated as specified in equation 10, and the sparsity of the matrix was estimated as exemplified in equation 11.

Currently, the parameter estimation in the gbggm package is handled by the YABS R package (Franco, 2023). YABS serves as a versatile tool for Bayesian Modeling, enabling users

Table 1

The true simulated partial correlation matrix

X	V01	V02	V03	V04	V05	V06	V07
V01	1.000	0.536	0.000	0.000	0.345	0.000	0.000
V02	0.536	1.000	0.340	0.515	0.000	0.000	0.000
V03	0.000	0.340	1.000	0.000	0.000	0.000	0.000
V04	0.000	0.515	0.000	1.000	0.000	0.306	0.000
V05	0.345	0.000	0.000	0.000	1.000	0.384	0.373
V06	0.000	0.000	0.000	0.306	0.384	1.000	0.340
V07	0.000	0.000	0.000	0.000	0.373	0.340	1.000

to implement Bayesian models exclusively using R functions. Comparable to LaplacesDemon (Hall et al., 2020) and fmcmc (Yon & Marjoram, 2019), YABS employs MCMC algorithms written in C++, resulting in accelerated calculations compared to these packages. Users can execute chains in parallel and select from four available samplers: Random-walk Metropolis, Metropolis-within Gibbs, Barker Proposal Metropolis (Livingstone & Zanella, 2022); or Oblique Hyperrectangle Slice Sampler (Thompson, 2011). Among these algorithms, Metropolis-within Gibbs and Oblique Hyperrectangle Slice Sampler are the most efficient, albeit slower. For smaller models, the Random-walk and Barker Proposal Metropolis samplers typically yield stable estimates with a reasonable number of samples after adequate adaptation steps.

For the estimation of the sparse GBGGM models, we utilized the Laplace Approximation method with Sampling-Importance Resampling (Kleppe & Skaug, 2012), also available in the YABS package. The model proposed by Williams (2021) was fitted with the BGGM R package version 2.0.4 (Williams & Mulder, 2020b). The glasso was fitted with the qgraph R package version 1.9.5 (Epskamp et al., 2012). To assess the models' performance, we computed seven performance statistics. Firstly, we calculated statistics based on the average error of the models: (i) Mean Absolute Error (MAE); (ii) Root-Mean-Squared Error (RMSE); (iii) MAE of the nodes that should be present; and (iv) MAE of the nodes that should be absent. Additionally, we assessed the accuracy of identifying present and absent edges through the following statistics: (i) Accuracy, reflecting the model's ability to correctly identify present and absent nodes; (ii) Sensitivity, indicating the model's ability to correctly identify nodes that should be present; and (iii) Specificity, representing the model's average ability to correctly identify nodes that should be absent. All materials required to reproduce this study are available in the supplementary materials

(<https://osf.io/6pk4n/>).

2.1 Results

The estimated partial correlations are presented in Table 2. For the Bayesian models, the shown values represent the MAP estimates. Due to the large number of values, we focus on specific observations. Firstly, the non-regularized partial correlations (shown in the "Partial.Poly" column) clearly and substantially overestimate the magnitude of the true partial correlations. Additionally, it appears that the glasso method outperforms the non-regularized partial correlations, although the estimated partial correlations remain larger than the true values overall. Notably, across the eight GBGGMs (from the "normal" to the "NEG" column) and the Williams model ("DW" column), the partial correlations—particularly those expected to be zero—are consistently lower than the true value.

To validate our intuitive observations from Table 2, we computed performance statistics, as displayed in Table 3. For MAE, RMSE, $er1$, and $er0$, values closer to zero are preferable. The laplace model exhibited the best average performance in terms of MAE, while the DW model demonstrated the best average performance in terms of RMSE. Interestingly, both the DW and laplace models performed equally well in terms of $er1$, representing the MAE of nodes expected to be present. However, the laplace model outperformed other models in terms of $er0$, representing the MAE of nodes expected to be absent.

The accuracy-based statistics (acc, sen, and spe) were derived from decisions regarding the presence or absence of a node, guided by the methodology of each model. In the case of glasso, absent nodes are determined by a tuning parameter established through an EBIC model selection method. For the GBGGMs, nodes are considered present if the sparsity parameter is estimated to be greater than 50%. Lastly, for the DW model, the presence of a

Table 2*True and estimated partial correlations*

Partial.Correlations	True	Partial.Poly	glasso	normal	laplace	logistic	cauchy	t	lomax	hypersec	NEG	DW
rho_01	0.536	0.816	0.680	0.506	0.525	0.675	0.559	0.631	0.580	0.541	0.329	0.334
rho_02	0.000	0.412	0.217	0.453	0.001	-0.332	0.000	-0.001	0.001	0.271	0.148	0.170
rho_03	0.000	-0.313	-0.432	0.000	-0.002	0.296	0.000	0.000	0.000	0.001	0.101	0.079
rho_04	0.345	0.647	0.342	0.000	0.001	0.016	0.000	0.000	-0.001	-0.252	0.202	0.067
rho_05	0.000	-0.991	0.008	-0.061	0.000	0.010	0.000	0.001	-0.001	0.000	0.144	0.041
rho_06	0.000	-0.603	-0.224	0.000	0.126	0.307	0.000	0.001	0.000	0.001	0.165	0.026
rho_07	0.340	0.191	0.162	0.000	0.460	0.492	0.258	0.326	0.314	0.000	0.189	0.246
rho_08	0.515	0.804	0.762	0.532	0.488	0.000	0.370	0.494	0.613	0.631	0.255	0.321
rho_09	0.000	-0.087	0.000	0.349	0.000	0.001	0.487	0.000	0.000	0.435	0.269	0.125
rho_10	0.000	0.887	0.360	0.000	0.000	0.193	0.000	0.001	-0.001	-0.001	0.206	0.076
rho_11	0.000	0.030	0.032	0.000	0.001	-0.001	0.000	0.001	0.000	0.000	0.232	0.043
rho_12	0.000	-0.737	-0.421	0.000	0.000	0.002	0.000	0.000	0.001	0.000	-0.016	0.049
rho_13	0.000	-0.961	-0.311	0.000	0.000	0.120	0.000	0.001	0.002	0.000	0.082	0.182



Table 2

True and estimated partial correlations

Partial.Correlations	True	Partial.Poly	glasso	normal	laplace	logistic	cauchy	t	lomax	hypersec	NEG	DW
rho_14	0.000	0.283	-0.171	0.226	0.000	0.007	0.000	0.301	0.000	0.356	0.006	0.044
rho_15	0.000	0.975	0.698	0.000	-0.001	0.176	0.000	-0.001	0.000	0.000	0.068	0.062
rho_16	0.000	-0.522	-0.236	0.000	-0.001	-0.001	0.000	0.272	0.207	0.000	0.064	0.184
rho_17	0.306	-0.440	-0.110	0.000	0.351	0.227	0.000	0.001	0.000	0.000	0.083	0.144
rho_18	0.000	0.570	0.456	0.000	0.000	0.186	0.000	0.000	0.000	0.000	0.098	0.105
rho_19	0.384	0.536	0.341	0.214	0.380	0.377	0.460	0.446	0.450	0.448	0.197	0.317
rho_20	0.373	0.998	0.662	0.555	0.827	0.380	0.491	0.635	0.771	0.000	0.184	0.215
rho_21	0.340	-0.487	-0.134	0.395	0.000	0.042	0.000	0.000	0.000	0.580	0.088	0.147



Table 3
Bias and accuracy indices for the estimated partial correlations

Model	MAE	RMSE	er1	er0	acc	sen	spe
True	NA	NA	NA	NA	NA	NA	NA
Partial Poly	0.511	0.595	0.421	0.567	NA	NA	NA
glasso	0.255	0.313	0.224	0.274	0.429	1.000	0.077
normal	0.121	0.192	0.180	0.084	0.667	0.625	0.692
laplace	0.070	0.150	0.168	0.010	0.857	0.750	0.923
logistic	0.150	0.209	0.190	0.126	0.571	0.750	0.462
cauchy	0.092	0.171	0.179	0.038	0.810	0.625	0.923
t	0.096	0.165	0.180	0.045	0.762	0.625	0.846
lomax	0.088	0.161	0.203	0.017	0.810	0.625	0.923
hyper-sec	0.148	0.236	0.255	0.082	0.714	0.625	0.769
NEG	0.153	0.170	0.201	0.123	0.619	0.000	1.000
DW	0.121	0.139	0.168	0.091	0.762	0.375	1.000

Note. MAE: Mean Absolute Error. RMSE: Root-Mean-Squared Error. er1: The MAE for the edges estimated to be present. er0: The MAE for the edges estimated to be absent. acc: Accuracy of the presence/absence of the nodes. sen: Sensitivity of the presence of the nodes. spe: Specificity of the absence of the nodes. NA: “Not A number”, indicating that statistic was not calculated for the given model.

node is determined by the 95% credible intervals for the partial correlations that do not encompass zero. These accuracy-based statistics vary from 0 to 1, with values closer to 1 considered as more adequate.

Using accuracy (acc) as a measure of overall performance, it becomes evident that the laplace model outperforms the others.

Concerning sensitivity (sen), which reflects the model's ability to accurately identify present edges, the glasso exhibited the best performance, while the NEG model showed the poorest sensitivity, recording zero. Regarding specificity, indicating the model's ability to accurately identify absent edges, both the NEG and DW models demonstrated optimal performance, achieving a specificity of one. Conversely, the glasso model displayed the weakest

Table 4*True and estimated thresholds*

	tau_01	tau_02	tau_03	tau_04	tau_05	tau_06	tau_07	MAE	RMSE
True	-1.500	-1.000	-0.500	0.000	0.500	1.000	1.500	NA	NA
normal	-1.347	-0.919	-0.508	-0.122	0.581	1.147	1.529	0.089	0.103
laplace	-1.307	-0.938	-0.494	-0.086	0.594	1.163	1.540	0.092	0.110
logistic	-1.322	-0.895	-0.468	-0.050	0.588	1.084	1.456	0.083	0.095
cauchy	-1.354	-0.918	-0.558	-0.057	0.566	1.113	1.438	0.083	0.089
t	-1.280	-0.980	-0.503	-0.057	0.615	1.105	1.544	0.081	0.106
lomax	-1.235	-0.927	-0.502	-0.113	0.575	1.158	1.529	0.102	0.131
hyper-sec	-1.320	-0.934	-0.528	-0.072	0.597	1.078	1.502	0.075	0.091
NEG	-1.300	-0.987	-0.509	-0.076	0.565	1.120	1.550	0.076	0.098

Note. MAE: Mean Absolute Error. RMSE: Root-Mean-Squared Error. NA: “Not A number”, indicating that statistic was not calculated for the given model.

specificity, with a score of 0.077.

Finally, Table 4 displays the thresholds estimated with the GBGGMs for the variables. In terms of overall performance statistics, we computed MAE and RMSE for each model. The hypersecant and NEG models exhibited the most favorable MAE values, while the Cauchy model demonstrated superior performance in terms of RMSE. However, overall, all models showed very similar performance in recovering the threshold values. Although the findings from Table 3 and 4 are not derived from extensive simulations, they nonetheless suggest promising prospects for generalized Bayesian Gaussian graphical models (GBGGMs).

3. ILLUSTRATION WITH AN EMPIRICAL EXAMPLE

In this empirical example, we used the second dataset from Faelens et al. (2019), available at

<https://osf.io/v7gch/>. This study investigated the negative association between Facebook use and mental health. Specifically, the authors explored the relationships between Facebook use, rumination, depressive, anxiety-, and stress-related symptoms, while considering key variables such as social comparison, contingent self-esteem, and global self-esteem. Data were collected from two independent samples: the first sample was used to generate exploratory hypotheses related to network centrality measures, while the second sample was employed in a preregistered replication study. The relationships between variables were estimated using the glasso method. Across both studies, social comparison and self-esteem emerged as central nodes in the network, linking social media use to indicators of psychopathology.

We conducted a reanalysis of the second dataset using the model proposed by Williams

(2021) and the laplace model from our GBGGM approach (i.e., equation 11), which showed promising performance in our toy simulation. This time, to fit the laplace model, we employed four different strategies: (i) sampling estimation with the Oblique Hyperrectangle Slice Sampler (OHSS) algorithm; (ii) Laplace Approximation with Sampling-Importance Resampling; (iii) sampling estimation with the No-U-Turn Sampler (NUTS); and (iv) Variational Inference. Strategies (i) and (ii) were implemented in the gbggm package with support from the YABS package, while strategies (iii) and (iv) were implemented using rstan (Gelman et al., 2015). Strategies (iii) and (iv) are also available in the supplementary materials, so the interested reader can check the Stan code to better understand how the GBGGMs are implemented. We employed these diverse strategies to examine whether, even when applying the same model, their results diverge. The model proposed by Williams (2021) was fitted with the BGGM R package version 2.0.4

(Williams & Mulder, 2020b).

Since there is no “true graph” for comparison, we evaluate the models by assessing their similarity in two key aspects: (i) the difference of the magnitude of the estimated partial correlations (after accounting for network sparsity), and (ii) their agreement on the presence or absence of edges. The first criterion is evaluated using the Mean Absolute Deviation (MAD) and Root-mean-squared deviation (RMSD) between the partial correlations estimated by each model/method. For the second criterion, we employ the normalized mutual information [NMI; Haghighat et al. (2011)] and the graph adjusted Rand index [GARI; Terada and Luxburg (2014)]. Lower MAD and RMSD values signify closer agreement in the magnitudes of the estimated partial correlation magnitudes across models/methods. Values of NMI and GARI closer to 1 indicate that the models/methods have estimated a more similar structure of absent/present edges. The exact values of the

Table 5

Differences between estimated partial correlations

	glasso	OHSS	LA	NUTS	VB	BGGM
glasso	0.000	0.041	0.042	0.037	0.059	0.043
OHSS	0.026	0.000	0.034	0.053	0.050	0.049
LA	0.026	0.014	0.000	0.050	0.044	0.049
NUTS	0.025	0.027	0.026	0.000	0.073	0.050
VB	0.037	0.024	0.019	0.042	0.000	0.065
BGGM	0.026	0.028	0.027	0.029	0.039	0.000

Note. glasso: glasso with EBIC model selection. LA: GBGGM estimated with Laplace Approximation with Sampling-Importance Resampling. OHSS: GBGGM estimated with the Oblique Hyperrectangle Slice Sampler (OHSS) algorithm. NUTS: GBGGM estimated with the No-U-Turn Sampler (NUTS). VB: GBGGM estimated with Variational Inference. Lower-diagonal values represent Mean Absolute Deviation (MAD) and upper-diagonal values represent Root-Mean-Squared Deviation (RMSD).

partial correlations are shown in the Appendix. Again, all materials required to reproduce this study are available in the supplementary materials (<https://osf.io/6pk4n/>).

3.1 Results

Table 5 presents the differences between the estimated partial correlations, with lower-diagonal values representing the Mean Absolute Deviation (MAD) and upper-diagonal values representing the Root-mean-squared deviation (RMSD). Overall, the models exhibit similar differences in their estimates. However, this uniformity can largely be attributed to the density of the networks, which refers to the proportion of edges estimated to be present. The glasso model demonstrates the highest density at 61.82%. In contrast, NUTS and BGGM exhibit densities of 43.64% and 47.23%, respectively. Meanwhile, OHSS, LA, and VB display densities of 32.73%, 30.91%, and 29.09%, respectively. These values indicate that while the glasso model estimates a dense network, the other methods estimate sparser networks.

Table 6 presents the agreement between the estimates of present and absent edges, with lower-diagonal values representing the Graph Adjusted Rand Index (GARI) and upper-diagonal values representing the Normalized Mutual Information (NMI). Notably, OHSS, LA, and VB achieved a GARI of approximately 90%, indicating substantial concordance regarding the inclusion of edges in the model. Particularly, the GBGGM estimated with NUTS and the BGGM demonstrated the highest agreement in terms of GARI compared to glasso. Moreover, when compared to BGGM, GBGGM estimates exhibited agreements exceeding 50%. Regarding NMI, NUTS displayed similar values with all models, including glasso. OHSS, LA, and VB also demonstrated considerable similarity in terms of NMI, aligning closely with BGGM. These findings suggest that the estimates of GBGGM remain consistent across various estimation methods, with discrepancies more apparent when NUTS is employed. However, it's important to note that these results could be

Table 6

Agreement between the density of the network

	glasso	OHSS	LA	NUTS	VB	BGGM
glasso	1.000	0.316	0.296	0.460	0.277	0.291
OHSS	0.170	1.000	0.877	0.438	0.791	0.514
LA	0.118	0.961	1.000	0.399	0.874	0.479
NUTS	0.415	0.698	0.660	1.000	0.363	0.317
VB	0.064	0.920	0.960	0.602	1.000	0.445
BGGM	0.274	0.668	0.627	0.585	0.585	1.000

Note. glasso: glasso with EBIC model selection. LA: GBGGM estimated with Laplace Approximation with Sampling-Importance Resampling. OHSS: GBGGM estimated with the Oblique Hyperrectangle Slice Sampler (OHSS) algorithm. NUTS: GBGGM estimated with the No-U-Turn Sampler (NUTS). VB: GBGGM estimated with Variational Inference. Lower-diagonal values represent Graph Adjusted Rand Index (GARI) and upper-diagonal values represent Normalized Mutual Information (NMI).

influenced by the hyperparameters utilized in each method, warranting further investigation.

4. FINAL REMARKS

We believe that the toy simulation and the empirical example show some promising results, as well as clear directions for future studies. Our main achievement was to demonstrate the feasibility of implementing models that offer flexibility in enabling various forms of regularization. Although our toy simulation featured eight different models, the potential exists to greatly expand this range by incorporating additional models based on alternative symmetric distributions. Moreover, more complex Bayesian regularization models, such as the horseshoe or hyperlasso, could further enrich the inference possibilities regarding the properties of the edges of a network. Another important achievement was showing that the at least the laplace model performs similarly with different estimation methods, indicating that one is free to choose which tool to use when implement GBGGMs.

For future studies, numerous possibilities emerge. Firstly, conducting a computer-intensive simulation study is imperative to validate the accuracy of the models proposed in our generalized approach. Because we did not conduct any systematic evaluation of the models, we refrain from advocating for the superiority of any specific GBGGM. However, we believe that our preliminary findings suggest a promising indication that these models can achieve performance at least comparable to that of the model proposed by Williams (2021). Additionally, the initial models implemented in this study likely represent the simplest models that can be implemented within our approach. Thus, they should be prioritized for extensive testing and evaluation.

Once the accuracy of the simplest models is firmly established, additional interesting aspects of our approach can be explored. For instance, the GBGGM approach facilitates the

estimation of moderated networks (Haslbeck et al., 2021) and, as a consequence, the comparison of group effects (Van Borkulo et al., 2022). For example, a moderated GBGGM with ridge regularization can be estimated for a set of standardized moderators M , in a model without an intercept, setting the prior for q_j as

$$q_j \sim N\left(M\beta_j, \frac{1}{\lambda}\right), \text{ for } j = i, \dots, p, \quad (12)$$

$$\lambda \sim \text{Cauchy}^+(0, 1)$$

where β_j is the regression parameter vector for the weight j of the Q matrix. This approach can also be extended to include vector generalized linear models (Yee, 2015), with the moderators influencing the amount of regularization. For the ridge regularization, with a set of standardized moderators M , this can be achieved with

$$q_j \sim N\left(M\beta_j, \frac{1}{\lambda} + M\kappa\right), \text{ for } j = i, \dots, p, \quad (13)$$

$$\lambda \sim \text{Cauchy}^+(0, 1)$$

where β_j is the same as before, and κ is the regression parameter vector for the effects of the moderators in the regularization. It is also somewhat easy to extend the models in equations 12 and 13 to include smooth functions of some predictor variables, resulting in vector generalized additive models (Yee, 2015). And, of course, the regression parameters can be regularized as well.

Another significant concern in psychometrics is the assessment of the dimensionality of measurement instruments. In network psychometrics, the commonly employed method is Exploratory Graph Analysis [EGA; Golino & Epskamp (2017)]. EGA entails a two-step process. First, a sparse partial correlations network is estimated, followed by dimensionality estimation using a community detection algorithm. Simulation studies (Cosemans et al., 2022) have demonstrated that the Louvain algorithm performs adequately in identifying the correct number of dimensions in data generated from latent factor models. Recently, Shi et al. (2023) introduced a Bayesian variant of EGA, employing the method proposed by Williams

(2021) for estimating the sparse partial correlations network instead of the conventional glasso method.

In our proposed generalized approach, a “fully Bayesian” implementation of EGA is conceivable. This could involve incorporating latent class cluster analysis into the model, utilizing a Dirichlet process to estimate the number of clusters (Müller et al., 2015). Additionally, an informative prior for modularity (Chen et al., 2014) of the full-sample sparse partial correlation matrix could be added, favoring models with higher modularity. This implies a prior assumption that network density within clusters is greater than between clusters (i.e., partial correlations should be higher between clusters rather than between). Such a model would yield a Bayesian EGA version with full posterior samples for both node communities in the model and the possible total number of communities. However, due to the considerable number of parameters involved, and the Bayesian non-parametric extension, an optimized sampling scheme would be most likely necessary.

One important unfavorable aspect of our approach is that it is difficult to think of informative priors in terms of the L matrix, and this task is even more cryptic if the parameters are thought of in terms of the elements of the C or Q matrices. One possible work-around for this issue is to include a second prior distribution regarding the correlation matrix implied by the model. This approach would allow one to use a Wishart prior, which is in the scale of the correlation matrix, to specify an informative prior for the implied covariance/correlation matrix. It is also possible to include individual priors for the elements of the correlation matrix implied by the model using normal priors for each. This “second prior” approach may seem unusual and some general purpose statistical software may not allow such calculations to be included in the model specification. However, this “second prior” can be seen as the posterior of a prior study (and it can actually be a posterior),

showing how, not only computationally, but also from a theoretical Bayesian perspective, coherent this approach is.

In conclusion, this study introduces the generalized approach to Bayesian GGMs. While our investigation has identified promising directions for future research, it represents only an initial step in a broader research agenda. Our proposed avenues for further study highlight the extensive opportunities awaiting exploration by researchers interested in Bayesian methodologies and network psychometrics. Furthermore, the flexibility and adaptability inherent in our approach hold significant potential for elucidating theoretical constructs across various domains within psychology. As this methodology continues to develop and undergo rigorous methodological scrutiny, it stands to make substantial contributions to our understanding of complex network structures and their implications.

5. REPRODUCIBILITY STATEMENT

The code and data that support the findings of this study, and that allows to completely reproduce our work, are openly available in Open Science Framework at <https://osf.io/6pk4n/>.

6. CONFLICTS OF INTEREST

The authors declare no competing interests.

7. AUTHOR CONTRIBUTIONS

The authors made the following contributions. V. R. F.: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing, Data analysis; G. W. F. B.: Writing - Review & Editing, Data analysis; M. J.: Writing - Review & Editing, Data analysis.

REFERENCES

- Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422), 669–679. <https://doi.org/10.1080/01621459.1993.1047632>

1

- Andersson, S. A., Madigan, D., & Perlman, M. D. (1997). On the markov equivalence of chain graphs, undirected graphs, and acyclic digraphs. *Scandinavian Journal of Statistics*, 24(1), 81–102. <https://doi.org/10.1111/1467-9469.00050>
- Archakov, I., & Hansen, P. R. (2021). A new parametrization of correlation matrices. *Econometrica*, 89(4), 1699–1715. <https://doi.org/10.1111/1467-9469.00050>
- Archakov, I., Hansen, P. R., & Luo, Y. (2024). A new method for generating random correlation matrices. *The Econometrics Journal*, 27(2), 188–212. <https://doi.org/10.1093/ectj/utad027>
- Arellano-Valle, R. B., & Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics*, 33(3), 561–574. <https://doi.org/10.1111/j.1467-9469.2006.00503.x>
- Bhat, C. R., & Mondal, A. (2021). *On the Almost Exact-Equivalence of the Radial and Spherical Unconstrained Cholesky-Based Parameterization Methods for Correlation Matrices*. Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin. https://www.caee.utexas.edu/prof/bhat/ABSTRACTS/Cholesky_parameterization.pdf
- Bindu, P., & Sangita, K. (2015). Double lomax distribution and its applications. *Statistica*, 75(3), 331–342. <https://doi.org/10.6092/issn.1973-2201/5190>
- Byrd, M., Nghiem, L. H., & McGee, M. (2021). Bayesian regularization of Gaussian graphical models with measurement error. *Computational Statistics & Data Analysis*, 156, 107085. <https://doi.org/10.1016/j.csda.2020.107085>
- Chen, M., Kuzmin, K., & Szymanski, B. K. (2014). Community detection via maximization of modularity and its variants. *IEEE Transactions on Computational Social Systems*, 7(1), 46–65. <https://doi.org/10.1109/TCSS.2014.2307458>
- Chen, Z., & Leng, C. (2015). Local linear estimation of covariance matrices via Cholesky decomposition. *Statistica Sinica*, 25(3), 1249–1263. <https://doi.org/10.5705/ss.2013.129>
- Cosemans, T., Rosseel, Y., & Gelper, S. (2022). Exploratory graph analysis for factor retention: Simulation results for continuous and binary data. *Educational and Psychological Measurement*, 82(5), 880–910. <https://doi.org/10.1177/00131644211059089>
- Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2019). Stability and variability of personality networks: A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, 136, 68–78. <https://doi.org/10.1016/j.paid.2017.06.011>
- D. Hoff, P. (2007). Extending the rank likelihood for semiparametric copula estimation. *The Annals of Applied Statistics*, 7(1), 265–283. <https://doi.org/10.1214/07-AOAS107>
- Déniz, E. G. (2021). Symmetric and asymmetric distributions: Theoretical developments and applications. *Symmetry*, 14(10), 2143. <https://doi.org/10.3390/sym14102143>
- Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika*, 85(1), 206–231. <https://doi.org/10.1007/s11336-020-09697-3>
- Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48, 1–18. <https://doi.org/10.18637/jss.v048.i04>
- Epskamp, S., Maris, G., Waldorp, L. J., & Borsboom, D. (2018). Network psychometrics. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development* (pp. 953–986). John Wiley & Sons Ltd. <https://doi.org/10.1002/9781118489772.ch30>
- Epskamp, S., Waldorp, L. J., Möttus, R., &

- Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, 53(4), 453–480.
<https://doi.org/10.1080/00273171.2018.1454823>
- Faelens, L., Hoorelbeke, K., Fried, E., De Raedt, R., & Koster, E. H. (2019). Negative influences of Facebook use through the lens of network analysis. *Computers in Human Behavior*, 96, 13–22.
<https://doi.org/10.1016/j.chb.2019.02.002>
 - Franco, V. R. (2023). *YABS: Yet another bayesian structure learning*. R package.
<https://github.com/vthorrf/YABS>
 - Franco, V. R., & Jimenez, M. (2023). *gbggm: Graphical bayesian generalized gaussian model*. R package.
<https://github.com/vthorrf/gbggm>
 - Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3), 432–441.
<https://doi.org/10.1093/biostatistics/kxm045>
 - Gelman, A., Lee, D., & Guo, J. (2015). Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics*, 40(5), 530–543.
<https://doi.org/10.3102/1076998615606113>
 - Ghosh, R. P., Mallick, B., & Pourahmadi, M. (2021). Bayesian estimation of correlation matrices of longitudinal data. *Bayesian Analysis*, 16(3), 1039–1058.
<https://doi.org/10.1214/20-BA1237>
 - Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS One*, 12(6), e0174035.
<https://doi.org/10.1371/journal.pone.0174035>
 - Gorbach, T., Lundquist, A., Luna, X. de, Nyberg, L., & Salami, A. (2020). A hierarchical Bayesian mixture model approach for analysis of resting-state functional brain connectivity: An alternative to thresholding. *Brain Connectivity*, 10(5), 202–211.
<https://doi.org/10.1089/brain.2020.0740>
 - Haghighat, M. B. A., Aghagolzadeh, A., & Seyedarabi, H. (2011). A non-reference image fusion metric based on mutual information of image features. *Computers & Electrical Engineering*, 37(5), 744–756.
<https://doi.org/10.1016/j.compeleceng.2011.07.012>
 - Hall, B., Hall, M., Statisticat, L., Brown, E., Hermanson, R., Charpentier, E., Heck, D., Laurent, S., Gronau, Q. F., & Singmann, H. (2020). *LaplacesDemon: Complete Environment for Bayesian Inference*. R package.
<https://cran.r-project.org/package=LaplacesDemon>
 - Hallquist, M. N., Wright, A. G., & Molenaar, P. C. (2021). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behavioral Research*, 56(2), 199–223.
<https://doi.org/10.1080/00273171.2019.1640103>
 - Hanada, M., & Matsuura, S. (2022). *MCMC from scratch: A practical introduction to markov chain monte carlo*. Springer Nature.
<https://doi.org/10.1007/978-981-19-2715-7>
 - Haslbeck, J. M., Borsboom, D., & Waldorp, L. J. (2021). Moderated network models. *Multivariate Behavioral Research*, 56(2), 256–287.
<https://doi.org/10.1080/00273171.2019.1677207>
 - Isvoranu, A.-M., & Epskamp, S. (2023). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods*, 28(4), 925–946.
<https://doi.org/10.1037/met0000439>
 - Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix. *Psychometrika*, 59(3), 381–389.
<https://doi.org/10.1007/BF02296131>
 - Kleppe, T. S., & Skaug, H. J. (2012). Fitting general stochastic volatility models using Laplace accelerated sequential importance sampling. *Computational Statistics & Data Analysis*, 56(11), 3105–3119.
<https://doi.org/10.1016/j.csda.2011.05.007>

- Krishnamoorthy, A., & Menon, D. (2013). Matrix inversion using Cholesky decomposition. In *2013 signal processing: Algorithms, architectures, arrangements, and applications (SPA)* (pp. 70-72). IEEE. <https://ieeexplore.ieee.org/abstract/document/6710599>
- Lange, K. L., Little, R. J., & Taylor, J. M. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, 84(408), 881-896. <https://doi.org/10.2307/2290063>
- Lauritzen, S. L. (1996). *Graphical models*. Oxford University Press. <https://doi.org/10.1093/oso/9780198522195.001.0001>
- Lee, S.-Y., & Poon, W.-Y. (1987). Two-step estimation of multivariate polychoric correlation. *Communications in Statistics-Theory and Methods*, 16(2), 307-320. <https://doi.org/10.1080/03610928708829368>
- Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9), 1989-2001. <https://doi.org/10.1016/j.jmva.2009.04.008>
- Li, Y., Craig, B. A., & Bhadra, A. (2019). The graphical horseshoe estimator for inverse covariance matrices. *Journal of Computational and Graphical Statistics*, 28(3), 747-757. <https://doi.org/10.1080/10618600.2019.1575744>
- Liu, C., & Martin, R. (2019). *An empirical G-wishart prior for sparse high-dimensional gaussian graphical models*. arXiv. <https://doi.org/10.48550/arXiv.1912.03807>
- Livingstone, S., & Zanella, G. (2022). The Barker proposal: Combining robustness and efficiency in gradient-based MCMC. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(2), 496-523. <https://doi.org/10.1111/rssb.12482>
- Marlin, B. M., & Murphy, K. P. (2009). Sparse Gaussian graphical models with unknown block structure. In *Proceedings of the 26th Annual International Conference on Machine Learning* (pp. 705-712). ICML. <https://doi.org/10.1145/1553374.1553465>
- Marsman, M., & Haslbeck, J. (2023). *Bayesian analysis of the ordinal Markov random field*. PsyArXiv. <https://doi.org/10.31234/osf.io/ukwrf>
- Marsman, M., Huth, K., Waldorp, L., & Ntzoufras, I. (2022). Objective Bayesian edge screening and structure selection for Ising networks. *Psychometrika*, 87(1), 47-82. <https://doi.org/10.1007/s11336-022-09848-8>
- Mohammadi, A., & Wit, E. C. (2015). Bayesian structure learning in sparse Gaussian graphical models. *Bayesian Analysis*, 10(1), 109-138. <https://doi.org/10.1214/14-BA889>
- Mohammadi, R., Massam, H., & Letac, G. (2021). Accelerating Bayesian structure learning in sparse Gaussian graphical models. *Journal of the American Statistical Association*, 118(542), 1345-1358. <https://doi.org/10.1080/01621459.2021.1996377>
- Moreno, J. L. (1934). *Who shall survive?: A new approach to the problem of human interrelations*. Nervous; mental disease publishing co. <https://doi.org/10.1037/10648-000>
- Müller, P., Quintana, F. A., Jara, A., & Hanson, T. (2015). *Bayesian nonparametric data analysis*. Springer.
- Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32(7), 685-694. <https://doi.org/10.1080/02664760500079464>
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443-460. <https://doi.org/10.1007/BF02296207>
- Park, J., Jin, I. H., & Schweinberger, M. (2022). Bayesian model selection for high-dimensional Ising models, with applications to educational data. *Computational Statistics & Data Analysis*, 165, 107325. <https://doi.org/10.1016/j.csda.2021.107325>
- Pearl, J. (2009). *Causality: Models, reasoning, and inference*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803161>
- Roverato, A. (2000). Cholesky decomposition

- of a hyper inverse wishart matrix. *Biometrika*, 87(1), 99–112. <https://doi.org/10.1093/biomet/87.1.99>
- Sagar, K., Banerjee, S., Datta, J., & Bhadra, A. (2024). Precision matrix estimation under the horseshoe-like prior-penalty dual. *Electronic Journal of Statistics*, 18(1), 1–46. <https://doi.org/10.1214/23-EJS2196>
 - Scheipl, F., & Kneib, T. (2009). Locally adaptive Bayesian P-splines with a Normal-Exponential-Gamma prior. *Computational Statistics & Data Analysis*, 53(10), 3533–3552. <https://doi.org/10.1016/j.csda.2009.03.009>
 - Scutari, M., & Silander, T. (2024). *bnlearn: Bayesian Network Structure Learning, Parameter Learning and Inference*. R package. <https://cran.r-project.org/package=bnlearn>
 - Shi, D., Christensen, A. P., Day, E., Golino, H., & Garrido, L. E. (2023). *A Bayesian approach for dimensionality assessment in psychological networks*. PsyArxiv. <https://doi.org/10.31234/osf.io/9rcev>
 - Sun, D., & Berger, J. O. (2007). Objective Bayesian analysis for the multivariate normal model. *Bayesian Statistics*, 8, 525–562. <https://doi.org/10.1093/oso/9780199214655.003.0020>
 - Talhouk, A., Doucet, A., & Murphy, K. (2012). Efficient bayesian inference for multivariate probit models with sparse inverse correlation matrices. *Journal of Computational and Graphical Statistics*, 21(3), 739–757. <https://doi.org/10.1080/10618600.2012.679239>
 - Terada, Y., & Luxburg, U. (2014). Local ordinal embedding. In *International Conference on Machine Learning* (pp. 847–855). PMLR. <https://proceedings.mlr.press/v32/terada14.html>
 - Thompson, M. (2011). *Slice sampling with multivariate steps* [Doctoral thesis, University of Toronto]. TSpace. <https://hdl.handle.net/1807/31955>
 - Van Borkulo, C. D., Bork, R. van, Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2022). Comparing network structures on three aspects: A permutation test. *Psychological Methods*, 28(6), 1273–1285. <https://doi.org/10.1037/met0000476>
 - Van Der Maas, H. L. J., Kan, K.-J., Marsman, M., & Stevenson, C. E. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence*, 5(2), 16. <https://doi.org/10.3390/jintelligence5020016>
 - Van Erp, S. (2020). A tutorial on Bayesian penalized regression with shrinkage priors for small sample sizes. In R. van de Schoot & M. Miočević (Eds.), *Small Sample Size Solutions* (pp. 71–84). Routledge. <https://doi.org/10.4324/9780429273872>
 - Van Erp, S., Oberski, D. L., & Mulder, J. (2019). Shrinkage priors for Bayesian penalized regression. *Journal of Mathematical Psychology*, 89, 31–50. <https://doi.org/10.1016/j.jmp.2018.12.004>
 - Verzelen, N. (2010). Adaptive estimation of covariance matrices via Cholesky decomposition. *Electronic Journal of Statistics*, 4, 1113–1150. <https://doi.org/10.1214/10-EJS580>
 - Wang, H. (2012). Bayesian graphical lasso models and efficient posterior computation. *Bayesian Analysis*, 7(4), 867–886. <https://doi.org/10.1214/12-BA729>
 - Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815478>
 - Williams, D. R. (2020). *Beyond lasso: A survey of nonconvex regularization in gaussian graphical models*. PsyArxiv. <https://doi.org/10.31234/osf.io/ad57p>
 - Williams, D. R. (2021). Bayesian estimation for gaussian graphical models: Structure learning, predictability, and network comparisons. *Multivariate Behavioral Research*, 56(2), 336–352. <https://doi.org/10.1080/00273171.2021.1894412>
 - Williams, D. R., & Mulder, J. (2020a). Bayesian hypothesis testing for Gaussian graphical models: Conditional independence and order constraints. *Journal of Mathematical*

Psychology, 99, 102441.

<https://doi.org/10.1016/j.jmp.2020.102441>

- Williams, D. R., & Mulder, J. (2020b). BGGM: Bayesian Gaussian graphical models in R. *Journal of Open Source Software*, 5(51), 2111. <https://doi.org/10.21105/joss.02111>
- Ye, Q., Amini, A. A., & Zhou, Q. (2020). Optimizing regularized cholesky score for order-based learning of bayesian networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10), 3555–3572. <https://doi.org/10.1109/TPAMI.2020.2990820>
- Yee, T. W. (2015). *Vector generalized linear and additive models: With an implementation in R*. Springer. <https://doi.org/10.1007/978-1-4939-2818-7>
- Yon, G. G. V., & Marjoram, P. (2019). fmcmmc: A friendly MCMC framework. *Journal of Open Source Software*, 4(39), 1427. <https://doi.org/10.21105/joss.01427>

APPENDIX

Table 7

Regularized partial correlations estimated for the empirical example

glasso	OHSS	LA	NUTS	VB	BGGM
0.196	0.258	0.226	0.225	0.246	0.203
0.053	0.000	0.000	0.092	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.212	0.214	0.280	0.217	0.259	0.214
0.244	0.330	0.270	0.292	0.382	0.296
0.291	0.244	0.276	0.308	0.208	0.279
0.000	0.000	0.000	0.000	0.000	0.107
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.020	0.000	0.000	0.048	0.000	0.000
0.376	0.441	0.438	0.445	0.390	0.402
0.108	0.000	0.000	0.174	0.000	0.106
0.000	0.000	0.000	0.000	0.000	0.000
0.087	0.230	0.175	0.000	0.172	0.161
0.037	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000

Table 7
Regularized partial correlations estimated for the empirical example

glasso	OHSS	LA	NUTS	VB	BGGM
-0.138	-0.174	0.000	-0.102	0.000	-0.139
-0.147	-0.165	-0.160	-0.141	-0.205	-0.191
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
-0.318	-0.273	-0.397	-0.358	-0.411	-0.376
0.123	0.141	0.159	0.118	0.176	0.127
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
-0.036	0.000	0.000	0.000	0.000	-0.115
0.243	0.220	0.276	0.191	0.332	0.237
0.077	0.000	0.000	0.000	0.000	0.000
0.375	0.445	0.434	0.436	0.395	0.366
0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
-0.083	-0.161	-0.182	-0.158	-0.143	-0.183
0.011	0.000	0.000	0.000	0.000	0.101
0.000	0.000	0.000	0.000	0.000	-0.108

Table 7
Regularized partial correlations estimated for the empirical example

glasso	OHSS	LA	NUTS	VB	BGGM
0.000	0.000	0.000	0.000	0.000	0.000
0.062	0.000	0.000	0.000	0.000	0.000
0.303	0.337	0.316	0.322	0.409	0.317
0.016	0.000	0.000	0.000	0.000	0.000
0.025	0.000	0.000	0.092	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000
0.126	0.114	0.143	0.146	0.000	0.146
0.004	0.000	0.000	0.050	0.000	0.000
0.031	0.000	0.000	0.000	0.000	0.000
0.027	0.000	0.000	0.000	0.000	0.119
0.038	0.000	0.000	0.094	0.000	0.115
0.004	0.000	0.000	0.000	0.000	0.000
-0.028	0.000	0.000	-0.069	0.000	-0.104
0.432	0.482	0.480	0.491	0.435	0.456
0.293	0.288	0.275	0.362	0.176	0.330
0.172	0.212	0.179	0.112	0.364	0.178

Note. glasso: glasso with EBIC model selection. LA: GBGGM estimated with Laplace Approximation with Sampling-Importance Resampling. OHSS: GBGGM estimated with the Oblique Hyperrectangle Slice Sampler (OHSS) algorithm. NUTS: GBGGM estimated with the No-U-Turn Sampler (NUTS). VB: GBGGM estimated with Variational Inference.